Artificial Neural Network Approach to Competency-Based Training Using a Virtual Reality Neurosurgical Simulation

Aiden Reich, MSc
Nykan Mirchi, MSc
Recai Yilmaz, MD
Nicole Ledwos, MSc
Vincent Bissonnette, MD, MSc
Dan Huy Tran, MSc
Alexander Winkler-Schwartz, MD
Bekir Karlik, PhD
Rolando F. Del Maestro, MD, PhD

Neurosurgical Simulation and Artificial Intelligence Learning Centre, Department of Neurology & Neurosurgery, Montreal Neurological Institute, McGill University, Montreal, Quebec, Canada

Correspondence:
Aiden Reich, MSc,
Neurosurgical Simulation and Artificial Intelligence Learning Centre,
Department of Neurology & Neurosurgery,
Montreal Neurological Institute,
McGill University,
3801 University St, Room E2.89,
Montreal, Quebec H3A 2B4, Canada.
Email: aiden.reich@mail.mcgill.ca
Twitter: @aidenreich

Received, March 15, 2021.
Accepted, November 8, 2021.
Published Online, May 11, 2022.

BACKGROUND: The methodology of assessment and training of surgical skills is evolving to deal with the emergence of competency-based training. Artificial neural networks (ANNs), a branch of artificial intelligence, can use newly generated metrics not only for assessment performance but also to quantitate individual metric importance and provide new insights into surgical expertise.

OBJECTIVE: To outline the educational utility of using an ANN in the assessment and quantitation of surgical expertise. A virtual reality vertebral osteophyte removal during a simulated surgical spine procedure is used as a model to outline this methodology.

METHODS: Twenty-one participants performed a simulated anterior cervical diskectomy and fusion on the Sim-Ortho virtual reality simulator. Participants were divided into 3 groups, including 9 postresidents, 5 senior residents, and 7 junior residents. Data were retrieved from the osteophyte removal component of the scenario, which involved using a simulated burr. The data were manipulated to initially generate 83 performance metrics spanning 3 categories (safety, efficiency, and motion) of which only the most relevant metrics were used to train and test the ANN.

RESULTS: The ANN model was trained on 6 safety metrics to a testing accuracy of 83.3%. The contributions of these performance metrics to expertise were revealed through connection weight products and outlined 2 identifiable learning patterns of technical skills.

CONCLUSION: This study outlines the potential utility of ANNs which allows a deeper understanding of the composites of surgical expertise and may contribute to the paradigm shift toward competency-based surgical training.

KEY WORDS: Anterior cervical diskectomy and fusion, Artificial neural network, Performance metrics, Surgical expertise, Surgical training, Virtual reality simulation

Operative Neurosurgery 23:31–39, 2022 https://doi.org/10.1227/ons.0000000000000173

Artificial neural networks (ANNs), a subset of artificial intelligence, function as interconnected nodes that communicate between each other and assign weights, which correspond to the algorithm's decision-making process.[1,2] These ANNs can uncover hidden performance patterns in large data sets by building connections and defining weights associated with each performance metric[3] and can then be exploited by competency-based training systems to develop new surgical education paradigms.[4,5]

ABBREVIATIONS: ACDF, anterior cervical diskectomy and fusion; **ANN,** artificial neural network; **CWP,** connection weight product; **VR,** virtual reality.

Supplemental digital content is available for this article at operativeneurosurgery-online.com.

Virtual reality (VR) anterior cervical diskectomy and fusion (ACDF) procedures allow trainees to develop a diversity of skills including knowledge, competence, and technical proficiency[6] and generate vast data sets of quantitative information relating to psychomotor skills.[1-3,7] Our group has explored VR simulation using performance metrics derived from the raw data that can be used by surgical educators to train individuals to improved levels of performance.[3,8-11] Studies focused on the educational utility of artificial intelligence techniques during surgical simulation have demonstrated that the classification of individuals into various groups is a valid method for assessing performance.[1,2,12-15] However, these approaches fail to outline the underlying reasons for classification and to quantify the relative importance of each metric used to train the model. An appreciation of the role that ANNs play in classifying performance, defining novel metrics, and quantitating the relative importance of each performance metric will enhance our understanding of the composites of surgical expertise and our ability to teach these composites.

The study objectives are (1) to introduce ANN methodology for assessing the composites of expertise in simulation-based training and (2) to outline the benefits of ANNs by using a VR spinal procedure model. This study reveals the potential of VR simulation combined with ANN to understand not only the essential composites of expertise and the contributions of each composite but also the interplay of specific composites to expert surgical performance.

METHODS

Participants

This study was a retrospective analysis of previous data that were used to develop face, content, and construct validity of the ACDF simulation on the Sim-Ortho VR platform.[3,16] Twenty-one participants performed the simulated ACDF on the Sim-Ortho VR platform. Participants were divided into 3 groups: 9 postresidents (4 practicing spinal neurosurgeons and orthopedic surgeons and 5 spine fellows), 5 senior residents, and 7 junior residents. Table 1 outlines participant demographic information and responses to a series of questions to assess experience and self-rated knowledge (on a 5-point Likert scale) related to the ACDF procedure which aided in predefining the 3 groups. All participants signed a consent approved by the local hospital research ethics board.

VR Simulator and Simulated Scenario

This study used the previously described ACDF simulation on the Sim-Ortho VR platform (OSSimTech) (Figure 1).[3,16] The simulated ACDF procedure was deconstructed into 4 components: disk annulus incision, diskectomy, vertebral osteophyte removal, and posterior longitudinal ligament excision (Video). The vertebral osteophyte removal portion that uses a burr had high face and construct validity in our previous studies, and this simulation was chosen to assess our objectives.[16]

Integration of ANN Within VR Simulation

The process of integrating an ANN to a VR simulator is illustrated in Figure 2. Raw data including information on position and angle of surgical instruments, forces applied, and tissue volume removed were constantly recorded and then manipulated and transformed into performance metrics. Initially, 83 metrics, encompassing safety, efficiency, and motion elements of simulation performance, were generated using expert opinion, publication data,[7,17] and novel metrics derived from raw data. Twenty-eight metrics were removed for comprising zero values, and after using a forward sequential feature selection algorithm, 6 metrics remained. The data from the 21 participants were split into 2 sets: 70% used for training the ANN (15 participants: 6 postresidents, 4 seniors, and 5 juniors) and 30% used for testing (6 participants: 3 postresidents, 1 senior, and 2 juniors). The training set data were used to train the ANN in a supervised manner, and the testing set data were used to test the model (**Methods**, **Supplemental Digital Contents 1-3** for more detailed information related to the ANN methodology, http://links.lww.com/ONS/A626, http://links.lww.com/ONS/A627, and http://links.lww.com/ONS/A628). The relative importance of each performance metric for each group was determined using the connection weights algorithm, which, in brief, considers the interconnected weighting system of the ANN (**Methods**, **Supplemental Digital Content 4**, http://links.lww.com/ONS/A629 and **Figure**, **Supplemental Digital Content 5**, http://links.lww.com/ONS/A630). Moreover, with the literature focusing on competency and defining it as a multifaceted set of skills culminating in desirable outcomes,[18] assessments of performance should be determined through a combination of metrics. Moving past the standard statistical analysis that evaluates metrics individually, the ANN methodology

TABLE 1. Demographics Information Related to Each Group of Participants for This Study

	Junior residents	Senior residents	Postresidents
No. of individuals	7	5	9
Age (y)			
Mean ± SD	27.4 ± 1.4	30.6 ± 2.3	44.2 ± 13.2
Sex			
Male	5	4	9
Female	2	1	0

	Level of training			
Surgical specialty	PGY 1-3	PGY 4-6	Fellows	Consultants
Neurosurgery	3	3	2	2
Orthopedic surgery	4	2	3	2
Relevant experience (percentage of group)				
Previously used a surgical simulator	5 (71%)	4 (80%)	7 (78%)	
Performed ACDF solo in the past month	1 (14%)	1 (20%)	7 (78%)	
Median self-rating on a 5-point Likert scale (range)				
Surgical knowledge of ACDF	3.0 (1.0-3.0)	3.0 (3.0-4.0)	5.0 (4.0-5.0)	
Comfort level performing ACDF alone	1.0 (1.0-3.0)	3.0 (2.0-4.0)	5.0 (3.0-5.0)	

ACDF, anterior cervical diskectomy and fusion; PGY, postgraduate year.

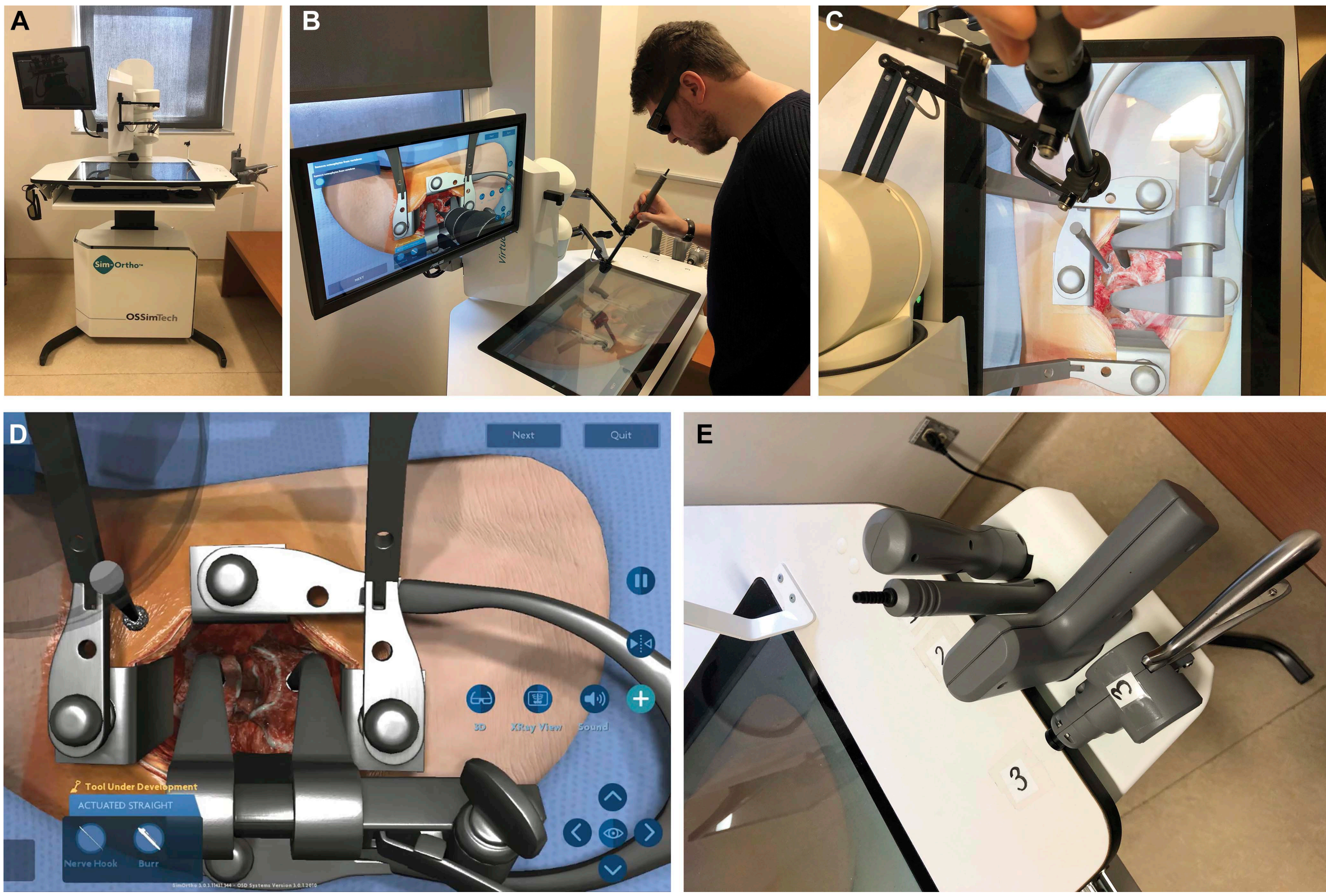

FIGURE 1. *The virtual reality simulator used for this study.* **A**, *The Sim-Ortho virtual reality simulator developed by OSSimTech.* **B**, *Utilization of 3D glasses and the haptic feedback arm—individual consented to the publication of his image.* **C**, *Participant's point of view on the Sim-Ortho Virtual Reality Simulator.* **D**, *Beginning of the osteophyte removal component of the anterior cervical diskectomy and fusion.* **E**, *Surgical instrument handles used on the Sim-Ortho virtual reality simulator. 3D, 3-dimensional.*

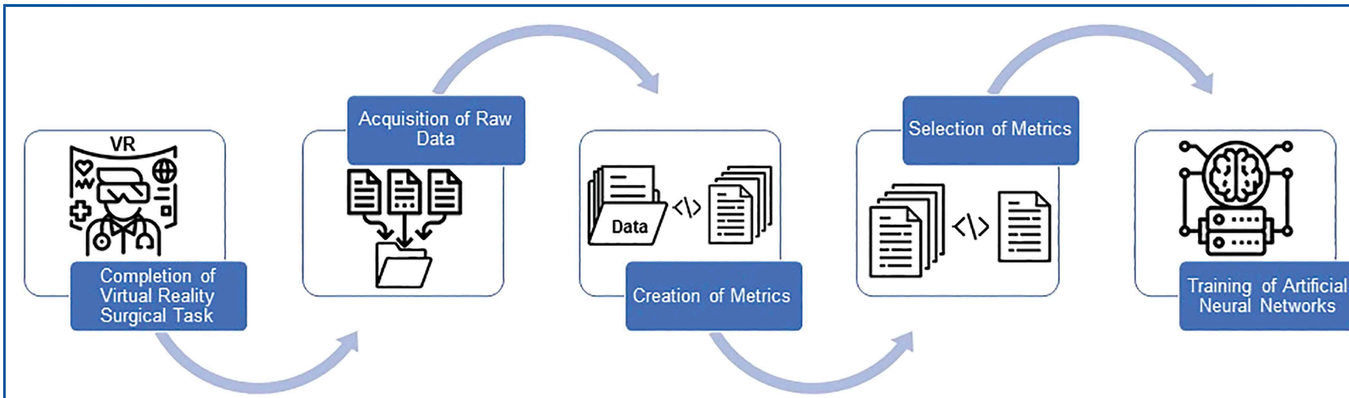

FIGURE 2. *Process of integrating artificial neural networks within virtual reality simulation. Participants perform the simulated scenario which generates a large data set of information related to their performance. The data are collected and transformed into quantifiable performance metrics. A feature selection is used to reduce the metrics to only the most relevant ones. These final metrics are used to train the artificial neural network.*

FIGURE 3. *Classification of the model.* **A**, *The confusion matrix of the training set correctly classified 12 of 15 participants reaching an 80% accuracy.* **B**, *The confusion matrix of testing set correctly classified 5 of 6 participants reaching an 83.3% accuracy.*

takes a holistic approach by analyzing the complete data of intimately interconnected metrics which together characterize surgical performance.

RESULTS

Performance Metrics and Classification of Surgical Performance

Six safety performance metrics were revealed for the osteophyte removal component of the ACDF scenario consistent with the predominance of safety metrics results from previous VR spine studies.[3,10] The confusion matrices separated into training (15 participants) and testing sets (6 participants), outlined in Figure 3, achieved overall accuracies of 80% and 83.3%, respectively. The training set misclassified 3 participants while the testing set misclassified only one. The design and optimization of the ANN are described in **Methods, Supplemental Digital Content 1-3**, http://links.lww.com/ONS/A626, http://links.lww.com/ONS/A627, and http://links.lww.com/ONS/A628.

Importance of Performance Metrics

The 6 metrics for postresidents, senior residents, and junior residents along with their corresponding connection weight products (CWPs) and relative importance are represented in Table 2. The magnitude of their CWPs ranged from +1.5 to −1.07 (Figure 4). Different metrics are important for each individual group, and these values are represented in Table 2.

Through the representation of the CWP in Figure 4, 2 main patterns were identifiable: continuous and discontinuous learning. The former involves a progressive and sequential change in which learning occurs in incremental stages or follows a certain prescribed order when comparing resident performance progress from junior to senior to postresident level of technical skills. The CWP associated with average force applied on the C5 vertebra (postresidents CWP = −1.07; senior residents CWP = −0.08; junior residents CWP = +0.68) illustrates this continuous learning pattern. These values signify an increased likelihood of being classified as a senior resident compared with a junior resident if the participant applied less average force on the C5 vertebrae while even lesser forces applied imply that the likelihood of being classified as postresident compared with a senior and junior resident is greater. The discontinuous learning pattern is coupled with a variable and nonsequential modification of skills learning as residents progress from junior to postresident performance level associated with senior resident incongruous performance. An example includes being classified as senior rather than postresident or junior resident if that participant has increased contact numbers on the spinal dura with an active burr (postresidents CWP = −0.68; senior residents CWP = +0.32; junior residents CWP = −0.92). The percentage of relative importance for the 6 metrics defined by the ANN provides new perspectives on critical aspects of expert performance during removal of vertebral osteophytes when using a burr (Table 2).

DISCUSSION

Summary

This study introduces the ANN methodology for assessing expertise in simulation-based training using the vertebral osteophyte removal component of the ACDF scenario. The utility of this ANN is outlined through the determination of the individual contributions to surgical performance of 6 metrics and associated learning patterns. The authors believe that an ANN, as the one used in this study, could be used in any medical or surgical procedure in which large data sets are available allowing one to assist with competency-based learning.

Utility of ANNs

Using this study's results and those from the diskectomy component of the procedure,[3] further analysis of the application

TABLE 2. List of Ranked Performance Metrics With Their Corresponding Weights and Relative Importance for Postresidents, Senior Residents, and Junior Residents

Expertise groups	Ranking of metrics	Performance metrics	Connection weight products	Relative importance (%)
Postresidents	1	Average force applied on right posterior longitudinal ligament	1.5	33.31
	2	Average force applied on C5 vertebra	−1.07	23.77
	3	Average force applied on left posterior longitudinal ligament	−0.71	15.78
	4	No. of contacts on the spinal dura with active burr	−0.68	15.06
	5	Average force applied on spinal dura	−0.30	6.72
	6	Average force applied on right vertebral artery	0.24	5.36
Senior residents	1	Average force applied on right posterior longitudinal ligament	−1.07	37.97
	2	Average force applied on left posterior longitudinal ligament	0.94	33.44
	3	No. of contacts on the spinal dura with active burr	0.32	11.27
	4	Average force applied on spinal dura	−0.30	10.56
	5	Average force applied on right vertebral artery	0.11	3.88
	6	Average force applied on C5 vertebra	−0.08	2.88
Junior residents	1	No. of contacts on the spinal dura with active burr	−0.92	27.54
	2	Average force applied on C5 vertebra	0.68	20.31
	3	Average force applied on right posterior longitudinal ligament	−0.64	19.04
	4	Average force applied on spinal dura	0.63	18.75
	5	Average force applied on right vertebral artery	0.33	9.95
	6	Average force applied on left posterior longitudinal ligament	−0.15	4.41

All 6 of these metrics were associated with safety.

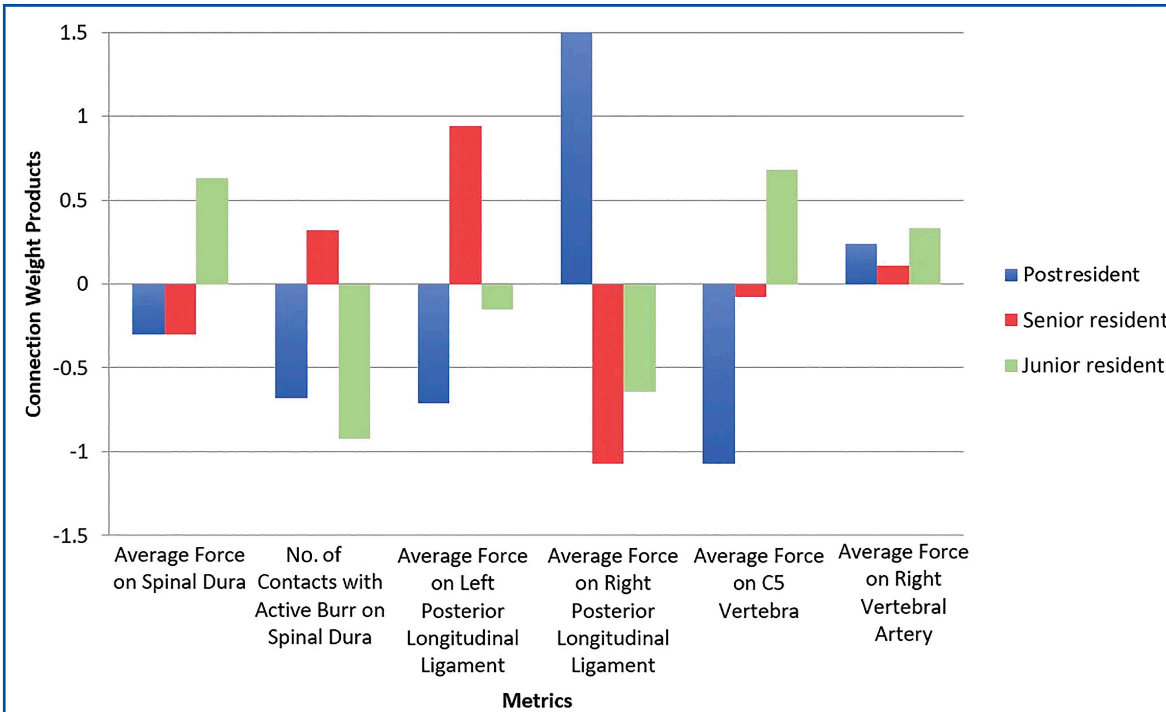

FIGURE 4. *Representation of performance metric connection weight products relative to each level of expertise. Each relative importance percentage corresponds to the magnitude of the connection weight product of a specific performance metric over the sum of all connection weight product magnitudes within the same performance level group of either postresident (blue), senior resident (red), or junior resident (green). Hence, the relative importance represents the influence of each metric on the classification of a participant into a specific group.*

TABLE 3. Comparison Between Diskectomy and Osteophyte Removal Components of the Simulated Anterior Cervical Diskectomy and Fusion

Aspects of comparison	Diskectomy	Osteophyte removal
No. of instruments used	3 (bone curette, pituitary rongeur, and disk rongeur)	1 (burr)
No. of metrics	16 metrics	6 metrics
Metrics categories	Safety, cognition, efficiency, and motion	Safety
Most important category of metrics	Safety (makes up more than 50% of the metrics)	Safety (makes up 100% of the metrics)
Testing accuracy	83.3%	83.3%
Connection weight product signs	Positive and negative	Positive and negative
Highest magnitude of connection weight product	5.24	1.5
Lowest magnitude of connection weight product	0.02	0.08
Hidden patterns	Continuous learning and discontinuous learning	Continuous learning and discontinuous learning

of the ANN in 2 different components of the same simulation procedure was possible (Table 3). The diskectomy component involved the participants choosing between 3 instruments while the osteophyte removal portion involved only using the burr. For complex operative procedures, the ANN methodology may outline both increased and more complicated metrics. In the diskectomy study, 16 metrics, involving safety, efficiency, motion, and cognition, were selected while only 6 safety metrics were selected for vertebral osteophyte removal component of the simulation. The predominance in safety metrics for both components of the simulation suggests that, despite differences in these 2 procedures, the ANN methodology regards operative safety as a critical composite of surgical expertise consistent with the previous data by our group.[10] As per previous studies,[3,13,19] the ANN methodology demonstrated its capability of differentiating individuals into surgical performance groups. With limited participant numbers in the training and the testing sets, lower classification accuracies such as 80% and 83.3% can occur.[20] An ANN model can be more robust by recruiting more individuals from various institutions and expanding both of the training and testing sets. Moreover, the CWP provides helpful insight into educators tasked with prioritizing specific metrics during training. For the 6 safety metrics identified for osteophyte removal, the relative importance percentage of these metrics is outlined in Table 2. These percentages allow educators to rank the performance metrics according to which contributes more toward the classification of a specific expertise level.

The outlining of these 6 safety metrics has allowed further precision and granularity in the classification of surgical expertise during the osteophyte removal component of the ACDF simulation procedure. Recent studies have identified the importance of ANN methodology in identifying important metrics in classifying expertise in the annulotomy[21] and diskectomy[3] components of the ACDF. The results of these studies allow the development of randomized control trials to assess the transferability of skill acquired during VR simulation to the patient in the operative room environment. These investigations will be critical to the development of new educational paradigms.

Educational Patterns of ANNs

The ability of ANNs to rank the significance of a specific performance metric during VR procedures allows surgical educators to investigate new concepts of teaching. Our results identified 2 patterns: the continuous and the discontinuous technical skills learning patterns (Figure 5). In the continuous pattern, decreasing average C5 vertebrae force application was associated with more advanced skills. Several studies involving VR from our group have also demonstrated that expertise is associated with decreasing force application consistent with these results.[2,3,22,23] The question does arise as to whether junior residents should be trained to the senior or postresident level of force application in this scenario to accelerate the acquisition of surgical expertise? The discontinuous learning pattern outlined involves the senior resident group performance being a nonsequential outlier compared with that of the junior and postresident groups. For example, this pattern was observed when the number of contacts of the active burr with the spinal dura metric was analyzed. It is reasonable to speculate that junior residents are hesitant to approach the spinal dura with the active burr while senior residents are more aggressive. With experience, postresidents may modulate this behavior with their predominant focus on safety. Considering this pattern, should junior residents be trained to the senior resident level of performance which is potentially associated with increased risk? Other patterns were also found in this study (Figure 4). An unusual pattern was that of the average force on the right posterior longitudinal ligament where the postresidents applied more force than both senior and junior residents. This may be a result of instrument positioning and hand ergonomics in relation to the ligament. One can speculate that postresidents were more aggressive when trying to completely remove osteophytes near the virtual patient's right posterior longitudinal ligament to decompress the adjacent cervical nerve. This action requires more wrist flexion, and the muscle activation associated with wrist flexion may lead to greater forces applied.[10] It would seem reasonable to develop studies to address the questions highlighted by the previous learning patterns including studies

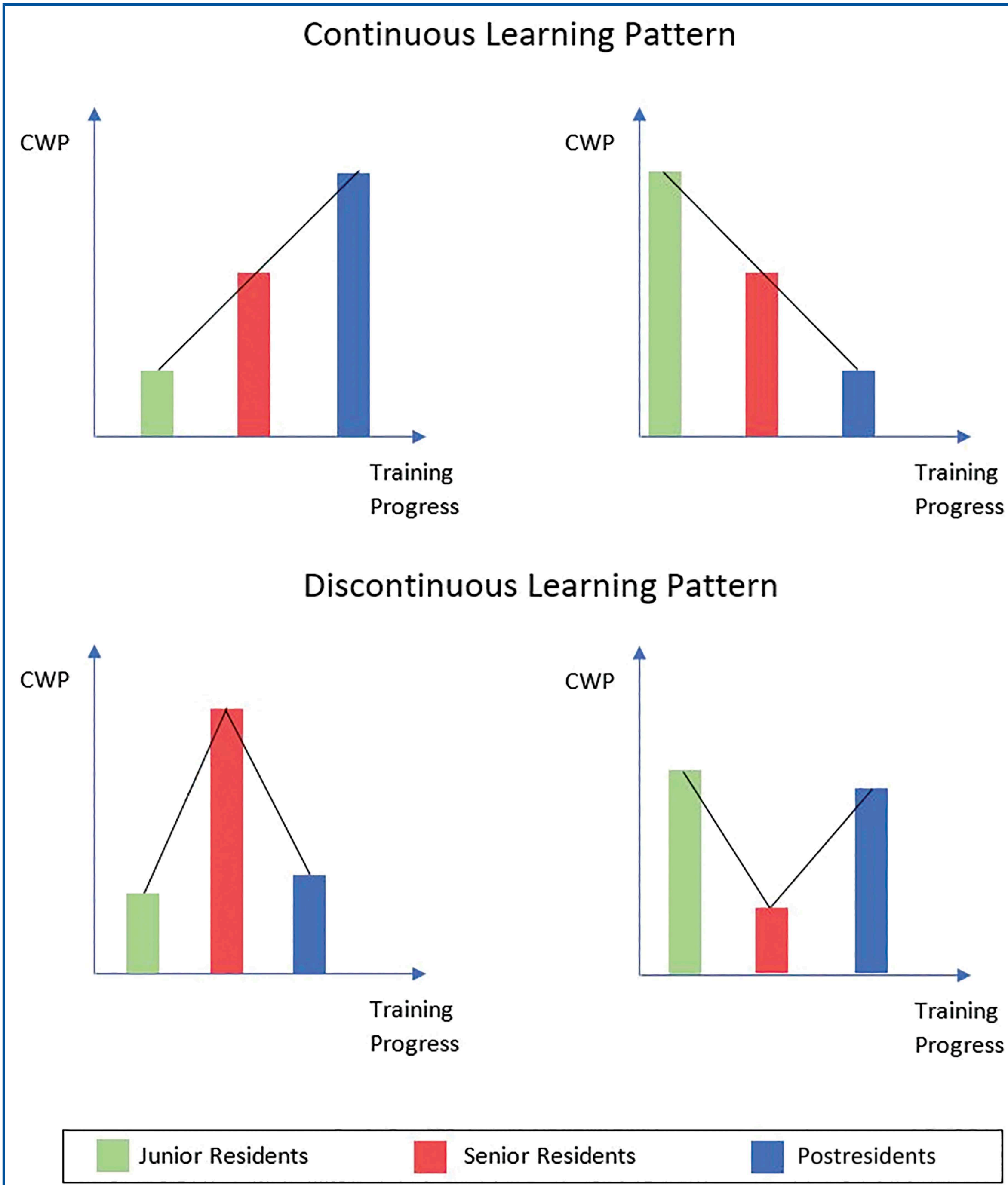

FIGURE 5. *Illustration of continuous and discontinuous learning patterns. With the x-axis representing the time variable training progress of an individual and the y-axis representing the connection weight product, the progressive and incremental learning curve associated with the continuous learning pattern and variable nonsequential learning curve of the discontinuous learning pattern are illustrated.*

involving a larger number of participants to investigate the less easily understood patterns. With these patterns, medical educators and trainees may adjust the focus of the training paradigms.

Surgical Education Platform Powered by ANNs

Reliable assessment tools that can effectively evaluate expertise are crucial for competency-based training.[24] Appropriate simulation technologies and the implementations of validated VR training scenarios can help to determine a trainee's baseline skill level and provide them with critical information to ensure quality feedback.[25,26] Simulation-based training may be automated by providing postoperative feedback to trainees using the information outputted by the ANN and, therefore, developed into intelligent tutoring systems.[27] This may decrease the reliance on expert instructors to give feedback or modulate the way these educators provide feedback.[27] With the benefits of a personalized and automated feedback system, a virtual operative assistant has been developed by our group,[28] and eventually, a virtual operative assistant could be developed using data obtained from ANN.

Limitations

Several limitations need to be considered. To limit task variables, participants were limited to only one specific burr size and speed, not consistent with burr options available to operating surgeons. Despite hand ergonomics proven to be an important factor in simulated operative procedure,[10] the Sim-Ortho platform was designed for right-handed use, suggesting modifications are warranted to assess the influence of ergonomics and bimanual performance. Training curricula and other factors differ depending on the surgical specialty and institution causing variable levels of expertise.[29,30] Because of the limited sample size of this study, the results cannot be extrapolated to larger population or groups originating from other institutions. Therefore, assessing the accuracy and generalizability of the utilization of ANN in competency-based training will require large prospective multi-institutional studies.

CONCLUSION

This study reveals the potential of VR surgical simulation combined with ANN to outline the important composites of surgical expertise, each composites' contribution, and their interplay. Through these utilities and learning patterns of technical skills revealed, both surgical educators and trainees can benefit from the ANN approach to competency-based training.

Funding

This work was supported by the AO Foundation (Davos, Switzerland) and the Franco Di Giovani Foundation.

Disclosures

The authors have no personal, financial, or institutional interest in any of the drugs, materials, or devices described in this article.

REFERENCES

1. Winkler-Schwartz A, Yilmaz R, Mirchi N, et al. Machine learning identification of surgical and operative factors associated with surgical expertise in virtual reality simulation. *JAMA Netw Open.* 2019;2(8):e198363.
2. Bissonnette V, Mirchi N, Ledwos N, Alsidieri G, Winkler-Schwartz A, Del Maestro RF. Artificial intelligence distinguishes surgical training levels in a virtual reality spinal task. *J Bone Joint Surg Am.* 2019;101(23):e127.
3. Mirchi N, Bissonnette V, Ledwos N, et al. Artificial neural networks to assess virtual reality anterior cervical discectomy performance. *Oper Neurosurg.* 2020;19(1):65-75.
4. Delorme S, Laroche D, DiRaddo R, Del Maestro RF. NeuroTouch: a physics-based virtual simulator for cranial microneurosurgery training. *Oper Neurosurg.* 2012;71(suppl 1):ONS32-ONS42.
5. Frank JR, Snell LS, Cate OT, et al. Competency-based medical education: theory to practice. *Med Teach.* 2010;32(8):638-645.
6. Ray WZ, Ganju A, Harrop JS, Hoh DJ. Developing an anterior cervical diskectomy and fusion simulator for neurosurgical resident training. *Neurosurgery.* 2013;73(suppl 1):S100-S106.
7. Sewell C, Morris D, Blevins NH, et al. Validating metrics for a mastoidectomy simulator. Paper presented at: MMVR2007.
8. Azarnoush H, Alzhrani G, Winkler-Schwartz A, et al. Neurosurgical virtual reality simulation metrics to assess psychomotor skills during brain tumor resection. *Int J Comput Assist Radiol Surg.* 2015;10(5):603-618.
9. Alotaibi FE, AlZhrani GA, Mullah MA, et al. Assessing bimanual performance in brain tumor resection with NeuroTouch, a virtual reality simulator. *Oper Neurosurg.* 2015;11(1):89-98.
10. Sawaya R, Alsideiri G, Bugdadi A, et al. Development of a performance model for virtual reality tumor resections. *J Neurosurg.* 2018;131(1):192-200.
11. Winkler-Schwartz A, Bissonnette V, Mirchi N, et al. Artificial intelligence in medical education: best practices using machine learning to assess surgical expertise in virtual reality simulation. *J Surg Educ.* 2019;76(6):1681-1690.
12. Watson RA. Use of a machine learning algorithm to classify expertise: analysis of hand motion patterns during a simulated surgical task. *Acad Med.* 2014;89(8):1163-1167.
13. Richstone L, Schwartz MJ, Seideman C, Cadeddu J, Marshall S, Kavoussi LR. Eye metrics as an objective assessment of surgical skill. *Ann Surg.* 2010;252(1):177-182.
14. Sewell C. *Automatic Performance Evaluation in Surgical Simulation*. PhD thesis. Stanford University; 2007.
15. Siyar S, Azarnoush H, Rashidi S, et al. Machine learning distinguishes neurosurgical skill levels in a virtual reality tumor resection task. Preprint. 2018. arXiv 181108159.
16. Ledwos N, Mirchi N, Bissonnette V, Winkler-Schwartz A, Yilmaz R, Del Maestro RF. Virtual reality anterior cervical discectomy and fusion simulation on the novel sim-ortho platform: validation studies. *Oper Neurosurg.* 2021;20(1):74-82.
17. Sewell C, Morris D, Blevins NH, et al. Providing metrics and performance feedback in a surgical simulator. *Computer Aided Surg.* 2008;13(2):63-81.
18. Gélinas-Phaneuf N, Del Maestro RF. Surgical expertise in neurosurgery: integrating theory into practice. *Neurosurgery.* 2013;73(suppl 1):S30-S38.
19. Yost MJ, Gardner J, Bell RM, et al. Predicting academic performance in surgical training. *J Surg Educ.* 2015;72(3):491-499.
20. Chang FM. Characteristics analysis for small data set learning and the comparison of classification methods. Paper presented at: 7th WSEAS International Conference on Artificial Intelligence, Knowledge Engineering and Data Bases, 2008.
21. Alkadri S, Ledwos N, Mirchi N, et al. Utilizing a multilayer perceptron artificial neural network to assess a virtual reality surgical procedure. *Comput Biol Med.* 2021;139:104770.
22. Azarnoush H, Siar S, Sawaya R, et al. The force pyramid: a spatial analysis of force application during virtual reality brain tumor resection. *J Neurosurg.* 2016;127(1):171-181.
23. Siyar S, Azarnoush H, Rashidi S, Del Maestro RF. Tremor assessment during virtual reality brain tumor resection. *J Surg Educ.* 2019;77(3):643-651.
24. Carraccio C, Wolfsthal SD, Englander R, Ferentz K, Martin C. Shifting paradigms: from Flexner to competencies. *Acad Med.* 2002;77(5):361-367.
25. Bing-You RG, Trowbridge RL. Why medical educators may be failing at feedback. *JAMA.* 2009;302(12):1330-1331.
26. Chan J, Pangal DJ, Cardinal T, et al. A systematic review of virtual reality for the assessment of technical skills in neurosurgery. *Neurosurg Focus.* 2021;51(2):E15.
27. Mirchi N, Ledwos N, Del Maestro RF. Intelligent tutoring systems: re-envisioning surgical education in response to COVID-19. *Can J Neurol Sci.* 2020;48(2):198-200.
28. Mirchi N, Bissonnette V, Yilmaz R, Ledwos N, Winkler-Schwartz A, Del Maestro RF. The virtual operative assistant: an explainable artificial intelligence tool for simulation-based training in surgery and medicine. *PLoS One.* 2020;15(2):e0229596.
29. Dvorak MF, Collins JB, Murnaghan L, et al. Confidence in spine training among senior neurosurgical and orthopedic residents. *Spine.* 2006;31(7):831-837.
30. Arnold PM, Brodke DS, Rampersaud YR. Differences between neurosurgeons and orthopedic surgeons in classifying cervical dislocation injuries and making

assessment and treatment decisions: a multicenter reliability study. *Am J Orthop (Belle Mead NJ).* 2009;38(10):E156-E161.

Acknowledgments

The authors wish to thank all the neurosurgeons, orthopedic surgeons, fellows and residents from the Montreal Neurological Institute and Hospital and the McGill Division of Orthopaedic Surgery who participated in this study. We also thank Drs Ghusn Ai Sideiri, Greg Berry, and Carlo Santaguida for their assistance with this study. We acknowledge OSSimTech^{TM} and the AO Foundation (Davos, Switzerland) for the development and use of the Sim-Ortho simulator platform.

***Supplemental digital content** is available for this article at operativeneurosurgery-online.com.*

Supplemental Digital Content 1. Methods. Feature selection.
Supplemental Digital Content 2. Methods. Design of the model.
Supplemental Digital Content 3. Methods. Optimization of the artificial neural network.
Supplemental Digital Content 4. Methods. Connection weights algorithm.
Supplemental Digital Content 5. Figure. Interconnectivity of the artificial neural network. The artificial neural network is composed of 3 layers. The input layer (left) comprises the performance metrics (X_x). Each performance metric is connected to every hidden layer (middle) perceptron (Y_x). The perceptrons of the hidden layer are themselves connected to each of the perceptrons (Z_x) in the output layer (right) representing each group of variant expertise level. Both the input hidden layer connections and the hidden output layer connections have weights (respectively, $W_{x,y}$ and $V_{x,y}$) attributed to them. These weights correspond to the sensitivity of each input on the algorithm's decision-making process.
VIDEO. The simulated anterior cervical diskectomy and fusion on the Sim-Ortho platform. A demonstration of the usability of the Sim-Ortho platform through a step-by-step explanation of the simulated neurosurgical procedure and its associated simulated surgical instruments.